1QBit

A Comparison of Text Sentiment and Market Sentiment

US Treasury 10-Year Note Futures
and Changes to Cash in Circulation using Sentiment Analysis and the CME Market Sentiment Meter

Pazinski Hong, Anish R. Verma

The CME Market Sentiment Meter (MSM) assigns market states from futures and options settlement data. There are four possible market states: Complacent, Balanced, Anxious, and Conflicted. This series of application notes provides detailed examples on how to use MSM by exploring economic events.

Executive Summary

The CME Market Sentiment Meter (MSM) calculates market sentiment states based on a novel mixture distribution, taking input from options and futures settlement data. We compare market sentiment from financial data to text sentiment from sentiment analysis as an indicator for market trends due to external events. Both types of sentiment were explored in a case study of the year 2020 about the US Treasury 10-Year Note futures (TYF). The year brought large fluctuations in the US economy due to the COVID-19 pandemic and other major events. The standard deviation of MSM mixture distribution has shifted earlier than the polarity score found through sentiment analysis. The movement in the MSM mixture distribution standard deviation preceded a price movement in the most active settlement price of the TYF.

1 Introduction

Constructing a reliable and consistent trading strategy has long been of interest for bankers, traders, and investors. Sophisticated and automated trading strategies generally perform better with more input of robust information, which can improve trading algorithms when looking to execute trade decisions.

Finding information that indicates a shift or unusual activity within the market is extremely valuable because it can be used as a trigger for trades that can reduce drawdown, or potentially increase return on investment. Some propose using social media text sentiment scores which can be calculated through a branch of natural language processing (NLP) called sentiment analysis. Generally, such a framework takes input from relevant social media and news feeds to calculate polarity scores, quantifying sentiment as positive, negative, or neutral. These scores may have some use in estimating price movements.

The CME Market Sentiment Meter (MSM) calculates market sentiment from actual market data, based on futures and options settlement data. The MSM outputs a sophisticated and diverse set of parameters, including four market sentiment states based on a novel risk-return distribution that takes into account multiple schools of thought in the market. The data is presented as a time series for easy implementation into existing workflows, and for rapid testing of trading strategies.

The year 2020 brought large fluctuations in the US economy due to the COVID-19 pandemic. The US Federal Reserve acted by adding more cash into circulation. Other major events include the UN Security Council Resolution 2532 regarding a global ceasefire and

1QBit

the US presidential election. Both text and market sentiment analysis were compared in a case study regarding the US Treasury 10-Year Note futures (TYF).

2 CME Market Sentiment States vs VADER Sentiment Analysis

From February to March 2020, indices like the S&P 500 saw roughly 30% declines, causing panic and concerns about the economy going forward. The US Federal Reserve acted by adding more cash into circulation, more colloquially referred to as "quantitative easing" or "printing money".

This had sharp effects on US Treasury 10-Year Note futures (TYF), as seen in Fig. 1. Treasury bonds are available in a variety of maturities. Generally, the relationships between the prices of these maturities give insight into how investors view the long-term development of the economy. Consequently, the prices of TYF are susceptible to events that affect government fiscal and monetary policies. In the year 2020, the shift in TYF reflected the sentiment of traders regarding the strength of the US economy and its ability to recover from the COVID-19 pandemic.
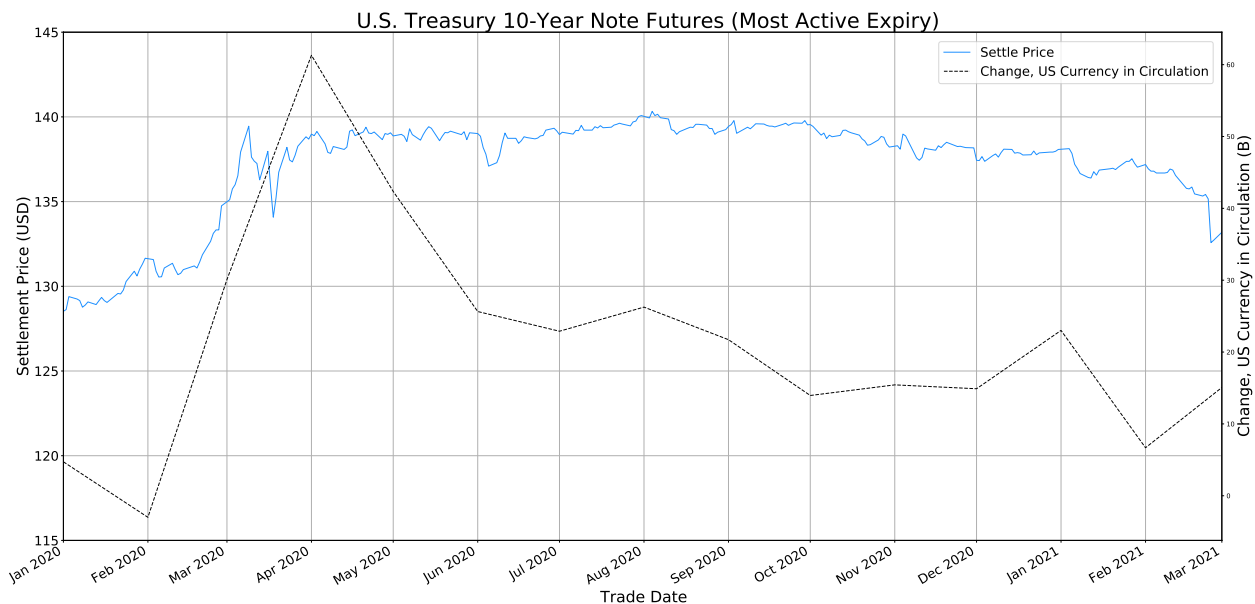

Figure 1: A comparison of the most active settlement price of TYF (solid blue) with the change in circulation (dotted black), January 2020–March 2021.

Surrounding this activity, there was much in the way of announcements and social media coverage. This motivates the study of sentiment analysis around the event to see if any insight can be gained from online text data.

For this case study, the Valence Aware Dictionary for Sentiment Reasoning (VADER) model was chosen to be used on raw speech data from the Federal Reserve. VADER is a

1QBit

rule-based NLP tool that is used for general sentiment analysis. Sentiment is quantified using the compound score returned by VADER. The score lies in the interval [-1, 1]. A common practice is to label sentiment as follows:

- Positive if a score is greater than or equal to 0.05;

- Neutral if the compound score lies between -0.05 and 0.05;

- Negative if the compound score is less than or equal to -0.05.

While VADER has shown its effectiveness on social media text data, it has been documented that the model can perform very poorly on text data consisting of frequent misspellings or sarcasm. However, speech from the federal reserve rarely consists of sarcasm or misspellings, making it an ideal candidate for VADER.

The results for the VADER sentiment analysis are plotted in Fig. 2. The sentiment analysis is juxtaposed with the most active settlement price for TYF, and the federal reserve data representing changes to cash in circulation. The green dots represent a positive sentiment score, and the orange dots represent a negative sentiment score. The purple dashed line represents monthly averaged sentiment scores.
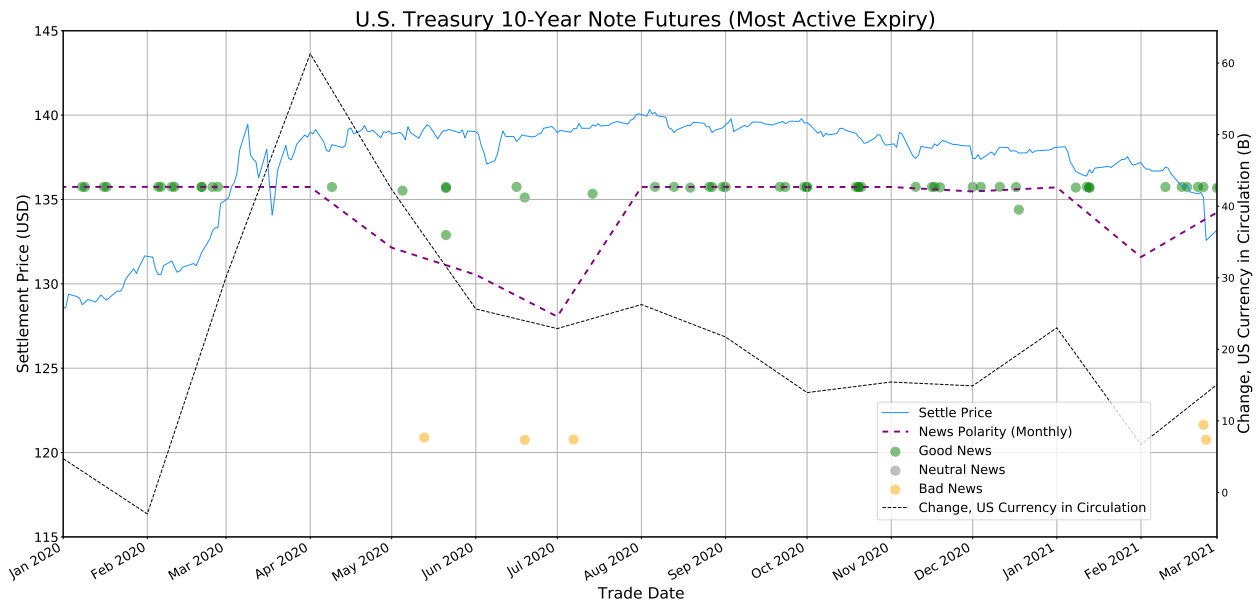

Figure 2: The results from the VADER sentiment analysis. A comparison of the most active settlement price of TYF (solid blue) with the change in circulation (dotted black), news polarity score from VADER (dashed purple), good news (green dot) and bad news (orange dot), from January 2020–March 2021. The polarity scores are first normalized then plotted to reflect this polarity, independent of the y-axis.

The polarity score drops after the increase in circulated cash. The cash flow increase happens after the rise in TYF, demonstrating that the polarity score is a lagging indicator for changes in TYF, and would not necessarily provide better results than looking at changes in the circulation of cash.
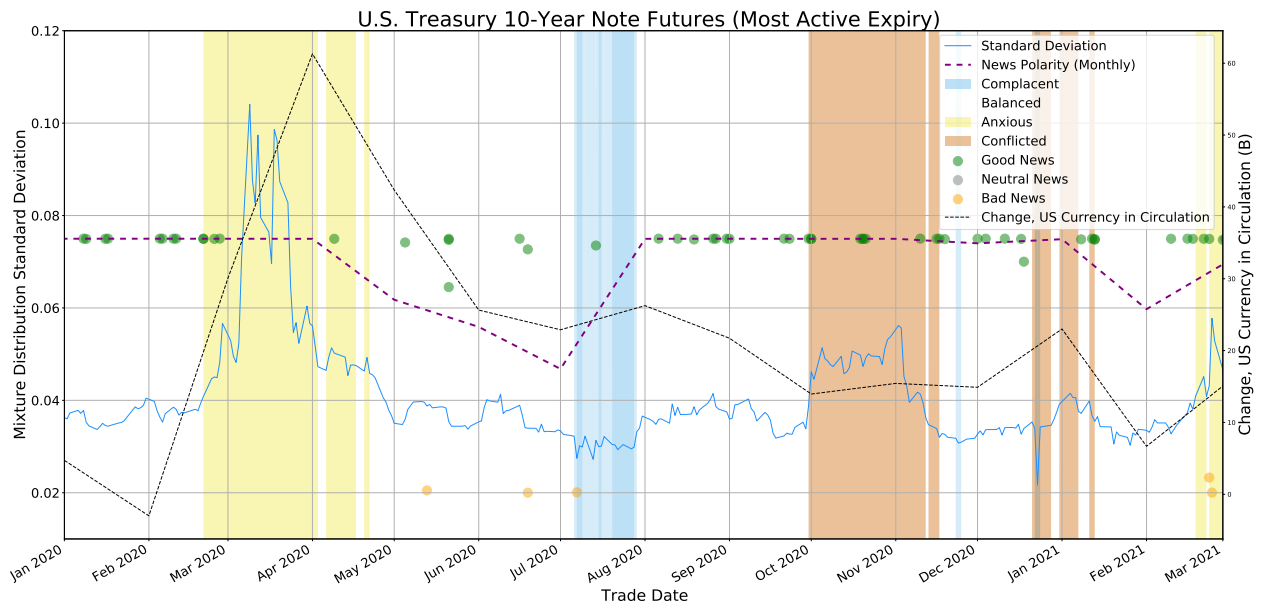
Figure 3: The MSM results overlaid on results from Fig. 2. The MSM mixture distribution standard deviation (solid blue) and the shaded regions correspond to one of the four market states: Complacent (blue), Balanced (white), Anxious (yellow) and Conflicted (red).

The CME Market Sentiment Meter (MSM) provides a measure of market sentiment based on options and futures data to construct a risk-return distribution. The distribution assumes that there can exist multiple schools of thought allowing for the classification of four market states. More detail is provided in the Appendix.

In Fig. 3 we append to the plot in Fig. 2 the MSM mixture distribution standard deviation. Of note, the mixture distribution standard deviation reacts to the shift in the cash in circulation. This is because it is based on trader decisions, inferred from options and futures settlement data, providing a quantitative gauge of market sentiment. The MSM showed that the market was in an Anxious state during the time of quantitative easing (March–April), resulting in a sharp spike in standard deviation, indicating a higher degree of risk for traders. The standard deviation quantifies risk: when it is larger, there is a higher chance of a larger price movement.

During the time of UN Resolution 2532, the MSM indicated the market was in a Complacent state with a decrease in standard deviation. For the US presidential election, the MSM indicated the market was in a Conflicted state which preceded a rapid increase followed by a decrease in standard deviation. In short, the MSM mixture distribution standard deviation shifted earlier than indicated by the polarity score found through the VADER analysis.

3   Conclusion

It was demonstrated that text sentiment is simply not enough to predict the shift of the market states, and in fact, the media sentiment is a reaction to the market. The CME Market Sentiment Meter can provide more useful information than VADER when applied to the raw speech data provided by the federal reserve.

Market sentiment analysis showed more forecasting power and insight into shifts or unusual activity within a market. This information can be used as triggers for trades that can reduce drawdown and increase return on investment.

1QBit

APPENDIX

A   Historical Market Analyses

This series of application notes describes notable historical events and times and how they affect market prices and trends. Each event considered is tailored to a specific product which is covered by the MSM product line. For each historical narrative, an analysis is provided making use of the data and features available through the MSM in order to gain an economic and financial insight into each situation.
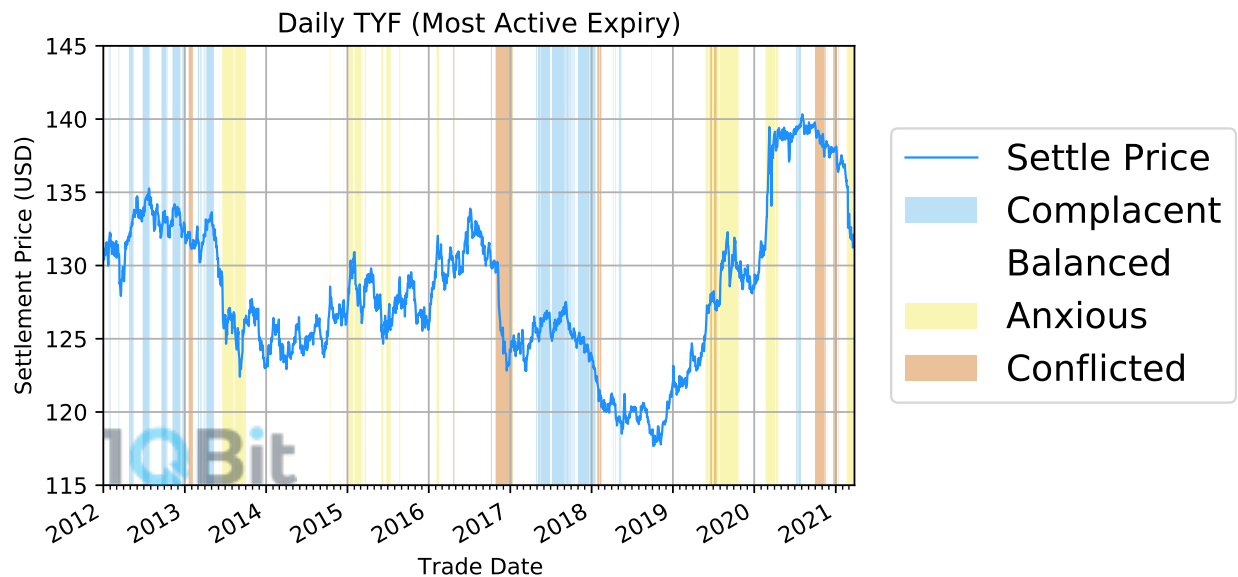

Figure 4: 2012-2019 Daily settlement price for TYF (most active expiry) (Blue line). The shading indicates the market state. Regions where the MSM is Complacent (Blue), Balanced (White), Anxious (Yellow), and Conflicted (Red) are highlighted.

B   CME Market Sentiment Meter (MSM) Market States

The MSM models risk by allowing there to be multiple "schools of thought" for price movement. There are four possible "market states": **Complacent**, **Balanced**, **Anxious**, and **Conflicted**. Graphical representations of the market states are found in Figs. 4-5. These are represented in a single Mixture Distribution. The Mixture Distribution represents the expected price movement over the next twelve months, so that its standard deviation can be directly compared to an annualized volatility. However, the Mixture Distribution may change daily, evolving over time as new information arises. This may indicate a large price move. The model works best for events in which the timing is known but the outcome is uncertain.
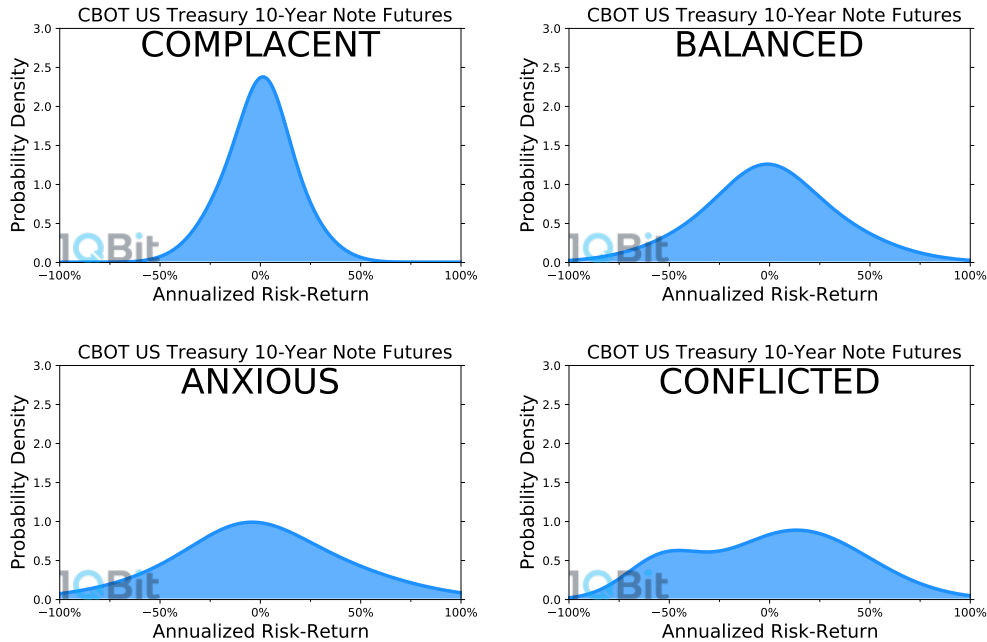
Figure 5: **MSM Market States.** Graphical representations of the risk-return curves for the four market states within the MSM: Complacent; Balanced; Anxious; and Conflicted states.

The **Complacent** market state is a "calm" state indicating that market participants have few concerns. Schools of thought do not differ significantly, resulting in a tall and narrow distribution. The narrow width of the curve is a direct result of a small standard deviation, and generally indicates that there is only a small chance of a large price move.

The **Balanced** market state is the most common state. This distribution has a larger standard deviation than the Complacent state, indicating a larger difference in the schools of thought when compared to the Complacent state.

In the **Anxious** market state, the schools of thought are diverging, and result in a much broader risk-return curve. Indeed, the differing schools of thought can also skew the distribution and move the mean off-centre, yielding information about the direction of a potential price move.

Most unique to the MSM model is the **Conflicted** market state. The defining feature of this, graphically, is the bi-modal nature of the distribution. For this situation, the schools of thought differ significantly and subsequently result in a large volatility.

C   CME DataMine Product Codes

Throughout the Historical Market Analyses, CME DataMine product codes are used to refer to the various products. For convenience, tabulated below is a list of CME DataMine product codes currently available through the MSM.

DataMine Code	Futures Product Name
C	CBOT Corn
CL	NYMEX WTI Crude Oil
EC	CME Euro FX (USD per EUR)
ES	CME E-Mini S&P 500
GC	COMEX Gold
NG	NYMEX Henry Hub Natural Gas
S	CBOT Soybean
TYF	CBOT US Treasury 10-Year Note

1QBit